

ABSTRACT

In the method and system for automated data diagnosis, a relational data base is formed of attributes characterizing an entity. The framework computes optimal diagnosis for a diagnostic inference problem which separates the desirable attributes from the undesirable attributes. The user of the system is permitted to specify certain parameters based on which the system computes a set of optimal or near optimal association rules between the attributes of a specific process, product, or other entity. The system and method of the present invention considers the simplicity of conditions in addition to support and confidence when ordering them, contributes the notion of tight conditions and semi-equivalence to remove redundant rules, uses the concept of the top fringes which allows near optimal conditions to be found, in addition to optimal conditions. Numeric as well as non-numeric attributes can be analyzed.